An Advanced Collaborative Environment to Enhance Magnetic Fusion Research

D.P. Schissel for the National Fusion Collaboratory Project
General Atomics, P.O. Box 85608, San Diego, California 92186-5608
E-mail: schissel@fusion.gat.com

Abstract

The SciDAC funded National Fusion Collaboratory Project has successfully started to deploy a production computational Grid for the United States fusion community. The MDSplus client/server data system used in fusion research has been outfitted with authentication and authorization capabilities by interfacing with the Globus and Akenti toolkits. A remote computation service using the TRANSP code at PPPL has been created and demonstrated to the user community for feedback on the general design. Visualization work has involved the creation of advanced 3D visualizations using SCIRun that retrieves data from MDSplus. Remote collaborative visualization has been demonstrated between tiled display walls. A prototype Mini–Access Grid node has been created to investigate the feasibility of reduced functionality and lower unit cost resulting in a larger number of deployed nodes

1. Introduction

The National Fusion Collaboratory project [1] was created to advance scientific understanding and innovation in magnetic fusion research by enabling more efficient use of existing experimental facilities through more effective integration of experiment, theory, and modeling. Fusion, the power source of the stars, has been the subject of international experimental and theoretical research since the late 1950's.

The goals of the Collaboratory are to (1) Create transparent and secure access to local/remote computation, visualization, and data servers, (2) Develop collaborative visualization that allows interactive sharing of graphical images among control room display devices, meeting room displays, and with offices over a wide area network, (3) Enable real–time access to high–powered remote computational services allowing such capabilities as between pulse analysis of experimental data and advanced scientific simulations. The computer science research necessary to create the Collaboratory is centered on three main activities: security, remote and distributed computing, and scientific visualization. This paper outlines progress to date on deploying the collaboratory prototype

2. Security and remote computing

The majority of the U.S. fusion community has adopted a common data management system, MDSplus, and a common relational database code-run–management schema implemented in Microsoft SQL Server for experimental and simulation data. The sharing of data, codes, and visualization tools as network services requires a system for protecting these valuable resources against unauthorized use. The Collaboratory is exploring state-of–the–art authentication, authorization, and encryption technologies provided by the Globus Security Infrastructure and the Akenti authorization service to work in concert with MDSplus and SQL Server to provide the necessary secure data services [2].

The Grid Security Infrastructure (GSI) component of the Globus toolkit uses X.509 identity certificates to provide for single sign–on authentication and delegation of user credentials to application programs. The Meta Data Service (MDS) provides GSI–enabled access to information pertaining to the availability and state of Grid resources. The Globus Resource Allocation and Management (GRAM) service defines a protocol and API for secure remote submission, monitoring, and management of compute jobs. The Akenti authorization service provides access policy specification for multiple resources based on credentials that are contained in certificates.

Initial work involved creating a prototype Fusion Grid by grid–enabling the client/server communications used in MDSplus remote access services to use Globus I/O services. The widely used fusion research code EFIT, that performs a magnetic equilibrium reconstruction, was modified to read and write securely to MDSplus. A demonstration was then performed at SC 2001 where a secure computation was controlled from the SC 2001 show floor by running EFIT on a Linux cluster at ANL with data being read and written to an MDSplus server at LBNL.

The success of the prototype system resulted in the decision to construct the production Fusion Grid. MDSplus data servers at the three main fusion experiments (C–Mod, DIII–D, and NSTX) were outfitted with Globus for secure communication. An alpha version of Globus I/O services on Windows was installed for secure communication via MDSplus to an SQL database. A prototype of the Akenti–enabled GRAM job manager was developed to perform authorization on job creation. The generally used fusion code TRANSP, that performs power balance and simulation analysis, was outfitted with both Globus and Akenti capabilities and installed on a multi–node Linux cluster at PPPL. Demonstrations of TRANSP computation as a Fusion Grid service was performed at two Fusion science meetings in April 2002. The demonstration included enhancements to Globus

monitoring capabilities as required by the fusion community. The deployment of TRANSP in the Fusion Grid was highly successful due both to the relative ease with which it could be deployed and by the positive feedback received from the fusion scientific community.

3. Scientific visualization

The demand placed on visualization tools by the Collaboratory is intense due to both the highly collaborative nature of fusion research and the dramatic increase in data resulting from the enhanced computing capabilities. The visualization component of the Collaboratory is focusing on the development of a collaborative control room, collaborative meeting room, and enhanced visualization tools.

Enhanced visualization tools that create a significant increase in capability and efficiency are required for both experimental and simulation data. Initial work has focused on using the Utah SCIRun problem solving environment for 3D visualization of fusion data [3]. SCIRun has previously been used for 3D visualization in the medical community and additions have been made to read fusion data from MDSplus. Specifically, the 3D plasma simulation code NIMROD, that is now storing data in MDSplus is being used as sample data for SCIRun visualization. Additionally, very modest cost Linux workstations with powerful graphics cards have been specified for rapid deployment to the fusion community. Demonstrations of this early work were made at the April 2002 Sherwood fusion theory meeting with positive feedback from the scientific community on both performance and cost of deployment.

Sharing of visualizations amongst a geographically distributed research staff is fundamental to the concept of a collaborative fusion control room. The MDSplus system gives data access to the worldwide community yet the ability to share visualizations back to the control room is presently not feasible. Sharing visualizations between two tiled display walls and workstation to tiled display wall has been investigated with both VNC and DMX [4]. Tiled wall to tiled wall collaborative visualization demonstrations were made between PCS and NSTX and ANL and DIII–D using VNC. Shared visualization involved traditional simple fusion visualization codes as well as the new fusion specific SCIRun application. The demonstrations generated considerable excitement about future possibilities amongst the fusion scientists.

The Access Grid technology is in use at over 80 institutions worldwide to support remote and distributed meetings. The use of AG nodes in the fusion community is under investigation for both the collaborative control room and collaborative meeting room. One thrust of this work is the investigation of a mini–AG node. Such a node would have reduced cost and would therefore be affordable to the smallest of fusion research sites. With reduce cost comes reduced functionality and current work is investigating the minimum specifications for a reduce node.

4. Discussion

The initial six months of the Fusion Collaboratory Project has proved very fruitful both in the education of the fusion scientist as well as in providing feedback to the Collaboratory Team on the initial Grid computing and visualization design. The demonstrations given at two fusion science meetings in April mimicked the introduction of a new laboratory into the Fusion Grid. This exercise has illustrated the importance of close coordination with site–security individuals who manage firewalls with Grid computing security offered via Globus and X.509 certificates. The complete development of a U.S. Fusion Grid that encompasses all 80 research sites will only be possible with a smooth integration of these two security administrative domains. The Fusion Collaboratory Project will pursue this integration.

The Project will continue to deploy the Fusion Grid with the release of TRANSP computational services in time for the annual fusion science meeting in November 2002 (APS/DPP). The release of this full service will require enhanced monitoring capabilities, refined authorization control, and initial investigation into job priority enforcement to support real–time analysis for fusion experiments. Visualization work will continue to develop the SCIRun application for use with fusion data. The creation of a general MDSplus interface to this application will allow its use for a wide variety of fusion data stored in MDSplus that goes substantially beyond the initial implementation with NIMROD. The collaborative visualization capabilities of both VNC and DMX will be pursued and a final design decision will be made in time for a demonstration at the APS/DPP meeting.

5. Acknowledgment

This work was supported by the U.S. Department of Energy SciDAC initiative and at General Atomics under Grant DE-FC02-01ER25455.

6. References

[1] http://www.fusiongrid.org/
[2] K. Keahey, et al., to be published in Future Generation Computer System (2001).
[3] http://www.sci.utah.edu
[4] http://www.uk.research.att.com/vnc/